**Introduction**

Realistic outdoor learning environments poses novel challenges for navigation following agents:
- Unique combination of a high variety of objects
- The transient nature of objects

**Street Simulator**

We built a simulator for navigation following agents as follows:
- For 10 US cities, we manually selected 1km² regions.
- The grid of locations in regions are separated 20 meters apart and generated using Google’s Street View API.
- As visual input at each location for the agent, we use 8 images covering 360° point of view for the agent.
- For each region, we randomly sampled 1000 start and end points.
- We generated navigation instructions using Google’s Directions API.

**Experiments**

**Baseline model**: sequence to sequence navigation following agent from the literature.

**Evaluation**: success rate and oracle success rate

**Conclusion**: Outdoor scenes from the real-world pose a challenging scenario for navigation following agents. We need natural language descriptions of routes that are more directly connected to the visual environment.

**Conclusion**: The large gap between the oracle success rate and the success rate for the baseline model shows that identifying where and when to stop remains an essential aspect of navigation following tasks.